The Social Construction of Digital Justice

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Jalen Gregory

Spring 2021

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature Jalen Gregory Date 05/11/2021
Jalen Gregory

Approved JUNDONG LI Date 04/29/2021
Jundong Li, Department of Computer Science

Approved _____ Date _____
Travis Elliott, Department of Engineering and Society

Introduction

In the information age, there is nothing as ubiquitous as data, the processing and analysis of information that we love being refined for the purpose of rendering the uncertainty that we hate a thing of the past. Machine learning has made it possible to take massive amounts of information and distill them into assumptions, if not outright assertions, and this has made data more powerful than ever. The question to ask is whether the insights gained from these machine learning algorithms are truly reflective of the world we aim to simulate and predict, or if they are simply reinforcing the beliefs of the engineers that build and entities that fund their creation.

The goal of this Science, Technology, and Society (STS) Research Paper is to study how the use of machine learning algorithms in criminal justice came about, what their implementation means to the varying social classes that encounter them, and how that will determine their use going forward. In their current form, most implementations of these algorithms are closed off to the public or "black-box", meaning what is going on under the hood is only available to those writing the code, leaving everyone else to base their opinions of them on their output. AI Now Institute co-founder Meredith Whittaker worries that that closed systems cut off from public judgment are dangerous, and "went so far as to recommend that no public agencies responsible for such matters as criminal justice, health care, welfare, and education should use black box AI systems" (Rieland, 2018). On the other side of the spectrum are those like University of Virginia Law alum and Fourth Circuit judge James Wilkinson III, who defended the practice of predictive policing in the face of what his fellow judges considered a clear 4th amendment violation in the case of US v. Curry (2020). He lamented that it is those with the fewest resources who tend to suffer the most when judges step in to stop law enforcement from using every tool at their disposal. The spectrum of opinions on the use of algorithms makes the social construction of

technology an ideal way to study their use and implementation as they are open to widely different interpretations based on the vantage point they are observed from.

Background

Algorithms help run the everyday world, even when people are not aware of their presence. In its base form, an algorithm is essentially a set of instructions that allow a set of inputs to produce an expected output. A popular example given to students by teachers and professors is the algorithm of making a peanut butter and jelly sandwich. This process is used to illustrate to students the steps that they inherently skip or take for granted when approaching a solution to a problem, steps that a computer is not able to skip. For years, even the most complex algorithms required expert knowledge to both transform into inputs and to assess the meaning of the outputs. In his proposal for a federal administration to regulate algorithms, Andrew Tutt (2016) describes how complex algorithms historically were programmed with inputs largely tuned by humans. Examples of this kind include Google's PageRank, the algorithm used to determine the importance of a given web page in relation to others, and IBM's Deep Blue, created to play chess against world champions. Machine learning, the process by which algorithms refine themselves, has changed the goal of algorithmic programming from one of training a machine to be good at a task to one of training a machine to learn what it means to be good at a task and how to get better (Tutt, 2016). Handing the reins over to the algorithm itself makes it difficult if not outright impossible to trace the "thinking" that led to any given decision, obfuscating the reasoning behind what could be a crucial decision.

The use of data in policing is not new but it has seen novel uses if not an outright renaissance in the machine learning age. Named one of the best inventions of 2011 by Time

magazine, predictive (or precision) policing is “the usage of mathematical, predictive analytics, and other analytical techniques in law enforcement to identify potential criminal activity” (Predictive policing, 2020). A leading company in this area is PredPol, which uses software that assumes crimes follow an earthquake aftershock model, where regions that previously experienced crime will continue to experience that crime with some decay over time (Ensign, et. al, 2017). Criminal justice entities face many of the same dilemmas faced in the private sector caused by issues of supply and demand in any number of areas. When police department budgets are stretched too thin or they are simply overrun with work, decisions must be made as to where to allocate their officers and how to approach the citizens in those areas. Judges in districts with overcrowded prisons often look for alternative solutions for those who they believe are less likely to recommit crimes but may not have all the information at hand to make individualized decisions in every case. That solution has been offered to them by computer scientists who have created tools called risk assessments, which are used to predict “future dangerousness of individuals” and estimate “flight risk and assess his or her threat to public safety” (Peeters & Schuilenburg, 2018). These risk assessments all have different inputs based on stakeholders decisions made in their creation, but tend to include historical data such as criminal history as well as any information that a judge would be able to assess from a court appearance (age, gender, etc.). For those not so lucky as to be offered an alternative to incarceration, the judges are still cognizant of a public that expects them to be fair in their judgment and not deliver overly harsh punishments based on information not necessarily relevant to the facts of the case. Machine learning can be and is used to guide all these decisions, but by its very nature makes the decisions reached difficult to predict and understand.

Framework: Social Construction of Technology

Overview

The Social Construction of Technology is a framework that emphasizes the role of society in shaping technology as opposed to the idea that technology shapes society (Pinch & Bijker, 2012). Put simply, in looking back on the successes and failures of technological innovations of the past, SCOT informs us that those successes and failures are driven as much by the social climate in which they were introduced as much as by their technical utility. A key concept of the social construction of technology is the idea of interpretative flexibility, or the way that different vantage points can reveal different strengths and weaknesses of a product. In their example following the development of the modern bicycle, Pinch and Bijker show that while the product is looked at now as the most "logical", it was the push and pull of competing interest groups that led it to that conclusion as opposed to any one independent factor. Social groups, collectives sharing identical interpretations of a technology, form the basis by which interpretative flexibility is understood. As the social groups that have an interest in a technology are derived, their concerns can be better understood or grouped. From there, the product continues to be adjusted to please the most important social groups and to least burden others as it stabilizes into what becomes the dominant model of the time. This stabilization is intended to lead to a closure whereby the problems pointed out by the social groups have either been eliminated or redefined.

Application to Machine Learning Algorithms

As explained by Pinch and Bijker (2012), it is important to realize that "all knowledge and all knowledge claims are sought in the domain of the social world rather than the natural world". While it is largely understood that the social beliefs of the time determine the laws, it is less widely understood that technology is susceptible to this same process. Algorithms, presented

as infallible as math equations, can give the impression that they are unbiased, but "if biased data is used to train these predictive models, the models will reproduce and, in some cases, amplify those same biases" (Lum & Isaac, 2016). Knowing this, it becomes critical that a study of the relevant social groups is undertaken to ensure that all their concerns are met. Unfortunately, as often is the case, the creation of the algorithms is left in the hands of a select few who may or may not have those concerns in mind.

At the forefront of the groups with an interest in criminal justice algorithms are the criminal justice officials who pay for their use. In the wake of mass protests in the summer of 2020, New York Police Department (NYPD) commissioner Dermot Shea, partially in response to calls for budget cuts, stated the need for his department of "better utilizing data, intelligence and all the technology at our disposal…That means for the NYPD's part, we'll redouble our precision-policing efforts" (Heaven, 2020). He is not nearly alone in this belief. It was former Los Angeles Police Department (LAPD) commissioner Charlie Beck whose staunch support for the algorithms led to them becoming more widely accepted across the nation. It was his belief that they were invaluable as they shift "law enforcement from focusing on what happened to focusing on what will happen and how to effectively deploy resources in front of crime, thereby changing outcomes" (Pearsall, 2010). As stated before, other public officials find benefits in their use as well. For example, in 2014, judges in Virginia using risk assessment tools trained by machine learning to determine the likelihood of defendants committing future crimes sent half of their defendants to alternatives to prison, slowing prison population growth to 5% between 2005 and 2016 from 31% in the 90's (Angwin et. al, 2016).

Outside of the criminal justice system, reaction to the creation of these algorithms is less positive. Experts on artificial intelligence, such as those at the aforementioned AI Now Institute,

have come out and lambasted their use, criticizing the championing of these algorithms as little more than “tech-washing, where a veneer of objectivity covers mechanisms that perpetuate inequities in society” (Heaven, 2020). In a letter titled “Abolish the #TechToPrisonPipeline”, 2435 experts across different fields lent their signature for the rescinding of publication of a research article that claimed to be able to predict criminality based on facial recognition on the basis that attempts to predict crime through technology do far more harm than good to the societies they aim to protect (Coalition for Critical Technology, 2020). With technology becoming more esoteric and complex as each year passes, the difficulty of reaching consensus among experts can make it appear as if conclusions can never be drawn: this case stands against that idea. The other main social group left in the equation is the citizenry that fall under the watchful eye of criminal justice algorithms. Some believe that they have nothing to hide, while others subscribe to the Edward Snowden belief that giving up rights because one has nothing to hide is akin to giving up freedom of speech because one has nothing to say. Many are not even aware of the implementation of software in the criminal justice system to the extent that it already has been. An LA resident, Hamid Khan, was so alarmed by his discovery of these algorithms that he has dedicated a large part of his life to fighting their adoption and raising awareness to fellow citizens. He demanded an audit of the algorithm used by PredPol but was rejected due to being told that the system was too complicated to determine pathways that led to the outputs given (Heaven, 2020). With the groups divided so fundamentally on the use of the algorithms, the next step of bringing them together to a conclusion seems difficult to achieve.

In attempting to synthesize what the opposing groups believe are the pros and cons of these algorithms, studies have been done to study their implementation as well as their impact. In their paper “When Politicization Stops Algorithms in Criminal Justice”, Konig and

Wenzelburger observe exactly the phenomenon that urged Pinch and Bijker to push forward the theory of SCOT in the first place, that being how little a role the effectiveness of the algorithms played in their final adoption (2021). Looking at 3 different states in their attempts to adopt risk assessments, they indicated that the single most important issue in the implementation of these tools was the way that they were framed to the public. In Alaska, after their adoption and a temporary increase in crime, an attack on risk assessments was launched on behalf of the bail industry, saying that they were making neighborhoods less safe than before. A similar situation occurred in Iowa where the assessments were curtailed on behalf a representative who believed that not enough research had been done to ensure the effects would be safe. In both cases, it was noted that the initial want to use the assessments was driven by subject matter experts on both sides of the criminal justice fence: public defenders and law enforcement agreed that the use of these assessments would ensure in a more safe and humane criminal justice system. These results show the imbalance of power between relevant social groups, and in a perfect example of SCOT, illustrate that technical efficacy is only one of many factors that impact the adoption of a technology.

The malleable nature of the law combined with the constantly transforming algorithms in use to assess them renders attempts at closure unlikely. The closure mechanism most reflective of the situation in my opinion is redefinition of the problem. As societal attitudes toward criminal justice solutions ebb and flow, the clashes between the laws we create to govern and the software we create to help deal with the consequences will shift. In assessing criminal justice algorithms in relation to the wider sociopolitical milieu, it could be said that they are emblematic of our current moment: technology leaned on to solve problems that the engineers that create them are unable to solve themselves. There is a technological hubris in their use that is representative of

our nation. Comparing the use of these algorithms to something from an episode of Black Mirror, Peeters and Schuilenburg argue that these tools represent further bureaucratization of our society. Arguing the algorithms themselves are bureaucratic, they note that "they structure human behaviour and decision-making in similar ways as bureaucratic organizations do – or perhaps even more so, given the reduction of judgement on individual cases to classification (Machine Justice, 2018). The designating of distinct social groups is made harder in this scenario because there are competing interests at stake: many believe that the criminal justice system needs immediate improvement while also maintaining a healthy skepticism of the entities that would be charged with the project of improving it. Whether or not a system that essentially classifies people as inputs given a set of vague criteria is useful, even if the result may be positive, is worth examining.

Conclusion

The debate over the adoption of machine learning algorithms in criminal justice can elucidate the priorities of our nation: would we prefer to work smarter instead of harder if it means less resources can be expended, and do we continue to shift responsibility from the hands of humans into the software of machines. As the law continues to adapt to protect an increasingly technological society, the use of software in legal settings is more than likely inevitable in some respect. It has yet to be proven whether these tools accomplish their stated goals, nevertheless they have been adopted and implemented in states and localities across the nation. To understand how this happened, one must understand the political and social contexts they were implemented in. Without a full effort in understanding the stakeholders involved in such a widespread technology, the effects can be more harmful than the danger they intend to prevent. Hailed by some experts as "maybe the biggest shift in the governance of security since the criminal justice

system began accepting social science and other expert evidence more than a century ago", the story of algorithmic justice will be an important one to keep tabs on going forward.

References

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine Bias. *ProPublica.* https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Coalition for Critical Technology. (2020, June 23). Abolish the #TechToPrisonPipeline. *Medium*. https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16

Ensign, D. et. al. (2017, Dec 22). Runaway Feedback Loops in Predictive Policing. *Proceedings of Machine Learning Research*, 81, 1-12. Retrieved May 10, 20221, from the arXiv dtabase.

Heaven, W. D. (2020, July 17). Predictive policing algorithms are racist. They need to be dismantled. *MIT Technology Review.* https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/

Konig, P. & Wenzelburger, G. (2021). When Politicization Stops Algorithms in Criminal Justice. (2021, Jan 22). *The British Journal of Criminology*. https://academic.oup.com/bjc/advance-article-abstract/doi/10.1093/bjc/azaa099/6105799

Lum, K. & Isaac, W. (2016). To predict and serve? *Significance, Vol. 13 (5),* 14-19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Pearsall, B. (2010). Predictive Policing: The Future of Law Enforcement? *National Institute of Justice Journal, Vol. 1* (266), 16-19. https://www.ncjrs.gov/pdffiles1/nij/230409.pdf

Peeters, R., & Schuilenburg, M. (2018). Machine justice: Governing security through the bureaucracy of algorithms. Information Polity, 23(3), 267-280. https://doi.org/10.3233/IP-180074

Pinch, T., & Bijker, W. (2012). The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. In Douglas D. (Author) & Bijker W., Hughes T., & Pinch T. (Eds.), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (pp. 11-44). Cambridge, Massachusetts; London, England: The MIT Press.

Predictive policing. (2020, November 29). In *Wikipedia*. https://en.wikipedia.org/wiki/Predictive_policing

Rieland, R. (2018, March 5). Artificial Intelligence is Now Used to Predict Crime. But Is It Biased? *Smithsonian Mag.* https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/

Tutt, A. (2016). An FDA for Algorithms. *Administrative Law Review, Vol. 69 (1),* 83-124.

United States v. Curry, No. 18-4233 (4th Cir. 2020). https://law.justia.com/cases/federal/appellate-courts/ca4/18-4233/18-4233-2020-07-16.html